Journal of Microbiological Methods 93 (2013) 42–48

Contents lists available at SciVerse ScienceDirect

Journal of Microbiological Methods

journal homepage: www.elsevier.com/locate/jmicmeth

Comparison of three molecular typing methods to assess genetic diversity for *Mycobacterium tuberculosis*

André Pitondo-Silva [a,*], Adolfo Carlos Barreto Santos [b], Keith A. Jolley [c], Clarice Queico Fujimura Leite [b], Ana Lúcia da Costa Darini [a]

[a] Departamento de Análises Clínicas, Toxicológicas e Bromatológicas, Faculdade de Ciências Farmacêuticas de Ribeirão Preto, Universidade de São Paulo, Brazil
[b] Departamento de Ciências Biológicas, Faculdade de Ciências Farmacêuticas, Universidade Estadual Paulista Júlio de Mesquita Filho, Brazil
[c] Department of Zoology, University of Oxford, Oxford, United Kingdom

ARTICLE INFO

Article history:
Received 24 November 2012
Received in revised form 29 January 2013
Accepted 31 January 2013
Available online 10 February 2013

Keywords:
M. tuberculosis
MIRU-VNTR
Spoligotyping
MLST

ABSTRACT

This study describes the comparison of three methods for genotyping of *Mycobacterium tuberculosis*, namely MIRU-VNTR (mycobacterial interspersed repetitive units-variable number of tandem repeats), spoligotyping and, for the first time, MLST (Multilocus Sequence Typing). In order to evaluate the discriminatory power of these methods, a total of 44 *M. tuberculosis* isolates obtained from sputum specimens of patients from Brazil were genotyped. Among the three methods, MLST showed the lowest discriminatory power compared to the other two techniques. MIRU-VNTR showed better discriminatory power when compared to spoligotyping, however, the combination of both methods provides the greatest level of discrimination and therefore this combination is the most useful genotyping tool to be applied to *M. tuberculosis* isolates.

1. Introduction

Tuberculosis (TB) remains as one of the most deadly infectious diseases worldwide and a leading public health problem. It is a chronic disease caused by the bacillus *Mycobacterium tuberculosis* that spreads from person to person through the air (Zaman, 2010). It is an infectious disease characterized by high morbidity and mortality in developing countries and in urban areas of developed countries (Sequera et al., 2008). Recently, the World Health Organization (WHO) announced, on the Global Tuberculosis Report 2012 that, in 2011, there were an estimated 8.7 million incident cases of TB globally, equivalent to 125 cases per 100,000 populations (WHO, 2012).

In the mid-1980s, the first integration of molecular methods to discriminate clinical isolates of *M. tuberculosis* was related. Previously established methods, such as comparative growth rates, colony morphology, susceptibility to select antibiotics, and phage typing are helpful, but present shortcomings due to their lack of discrimination, which restrains their application in TB epidemiology (Mathema et al., 2006).

Given the evolution of DNA-based approaches in recent decades, many genotypic methods have been used effectively in taxonomic and identification studies of a range of bacterial genera, including *M. tuberculosis*. Furthermore, whole-genome sequence data have opened up new insights into epidemiological surveillance (Lukinmaa et al., 2004). DNA fingerprinting of *M. tuberculosis* isolates is a helpful technique to establish the extension of recent transmission in a population and the probable risk factors for recent transmission, to identify earlier unsuspected transmission, to screen the transmission of drug-resistant strains, and to confirm laboratory cross contamination (Cowan et al., 2002).

Several alternative PCR-based methods have been developed in order to overcome these problems. Spoligotyping is often used as a secondary method for typing low-copy-number IS*6110* isolates, but it does little to improve strain differentiation in population-based studies (Allix-Béguec et al., 2008; Cronin et al., 2001). Spoligotyping is a technique related to DNA polymorphisms within direct repeat (DR) locus of *M. tuberculosis*, pointing out the presence or absence of a set of target spacer sequences between the DRs. The octonal pattern of spacers between the conserved DRs in the region was used to differentiate *M. tuberculosis* strains (Mendes et al., 2011; Sola et al., 2001).

Currently, genotyping methods aimed at the analysis of the variable number of tandem repeats (VNTR), using mycobacterial interspersed repetitive units (MIRU), present the best potential. This technique is based on the variability found in 12 specific loci interspersed throughout the mycobacterial genome (Supply et al., 2000). Lately, MIRU-VNTR genotyping approaches employing 15 or 24 loci (Alonso-Rodríguez et al., 2008; Supply et al., 2006) have been

* Corresponding author at: Departamento de Análises Clínicas, Toxicológicas e Bromatológicas, Faculdade de Ciências Farmacêuticas de Ribeirão Preto, USP, Av. do Café, s/n, Monte Alegre, Ribeirão Preto, SP, CEP: 14040-903, Brazil. Tel.: +55 16 3602 4290; fax: +55 16 3602 4725.
E-mail address: andre@pitondo.com.br (A. Pitondo-Silva).

0167-7012
http://dx.doi.org/10.1016/j.mimet.2013.01.020

evaluated and applied to molecular epidemiological typing in mycobacteria (Noguti et al., 2012).

Recently, the analysis of multiple housekeeping genes has become a broadly applied tool for the investigation of taxonomic relationships for many bacterial species. It is possible to attain an overall and consistent indication of genetic relationships among different organisms by using information obtained by the comparison and combination of multiple genes. The ad hoc committee for re-evaluation of the species classification has established that the sequencing of a minimum of five well-chosen housekeeping genes, universally distributed, present as single copies and located at distinct chromosomal loci, is a promising method for bacterial typing (Martens et al., 2008; Stackebrandt et al., 2002).

Multilocus sequence typing (MLST) is a technique for typing multiple loci using the DNA sequences of internal fragments of multiple housekeeping genes, which was proposed in 1998 as a portable, universal, and definitive method for characterizing bacteria (Maiden et al., 1998). Thereafter, MLST data have been increasingly employed in epidemiological investigations of various scales and in studies of population biology, pathogenicity, and evolution of different bacterial species (Maiden, 2006).

The purpose of this study was to compare the two main molecular typing methods for *M. tuberculosis*, namely MIRU-VNTR and spoligotyping, and also MLST for this species.

2. Material and methods

2.1. Bacterial isolates and patients

A total of 44 *M. tuberculosis* isolates were obtained from sputum specimens of patients over 18 years of age with pulmonary TB, between 2002 and 2007, from four different cities in Brazil, namely São Paulo, São Paulo State (SP) (n = 10), Araraquara, SP (n = 12), Dourados, Mato Grosso do Sul State (MS) (n = 20) and Campo Grande, MS (n = 2). *M. tuberculosis* H37Rv (ATCC 27294) strain was used as positive control in all experiments. Patients received medical care at the following reference health centers for TB treatment: Special Health Service of Araraquara, (SESA) located in Araraquara city, SP; Clemente Ferreira Institute, located in São Paulo city, SP; and Central Laboratory of Mato Grosso do Sul (LACEN-MS), located in Campo Grande city, MS. These reference centers were responsible for the diagnosis of tuberculosis (smear microscopy, culture and phenotypic identification tests). *M. tuberculosis* isolates from Dourados and Campo Grande were isolated from indigenous patients and the remaining isolates were obtained from outpatients.

Mycobacteria isolates identified as belonging to the *M. tuberculosis* complex by phenotypic methods were sent to the Mycobacteriology Laboratory (FCFAR-UNESP), São Paulo, Brazil, for molecular assays including polymerase chain reaction (PCR), identification of IS*6110* and genotyping. All isolates were inoculated in Löwenstein–Jensen (LJ) solid medium, incubated at 37 °C and identified as *M. tuberculosis* by PCR-IS*6110* (Leão et al., 2004; van Embden, et al., 1993). This study has been approved by the Research Ethics Committee of the Faculdade de Ciências Farmacêuticas de Ribeirão Preto, USP (reference 35/2008). As for indigenous samples, the approval was issued by the National Research Ethics Committee (CONEP, reference 869/2009).

2.2. MIRU-VNTR typing

Genomic DNA was extracted by thermolysis method, as previously described (Mazars et al., 2001). As for genotyping, the set of VNTRs, consisting of 12 MIRUs (MIRU loci 2, 4, 10, 16, 20, 23, 24, 26, 27, 31, 39, 40) was investigated as mentioned (Mazars et al., 2001; Supply et al., 2000, 2002). Each locus was amplified individually in a PTC-100 thermal cycler (MJ Research, Ramsey, Minnesota, USA) and the amplified products were resolved by 1.5% agarose gel electrophoresis into bands, which were stained with ethidium bromide (0.5 μg mL^{-1}) and revealed under UV light. The size of the PCR fragment for each locus was estimated by visual comparison with the molecular markers and the MIRU allele score was determined by the size, as described (Supply et al., 2002; Mazars et al., 2001). Results from each of the 12 loci were combined to create a 12-digit allelic profile.

2.3. Spoligotyping

Spoligotyping was performed in all *M. tuberculosis* isolates using the standard method to detect the presence or absence of 43 spacers (Kamerbeek et al., 1997). The general PCR procedure, including preparation of the sample, master mix and primers, was performed by the method given by Mendes et al. (2011). PCR products were hybridized to a spoligo-membrane (Pall Biosystems, Portsmouth, UK) containing 43 spacer oligonucleotides covalently linked therein. Spoligotype results were converted into octal code and in the SITVIT2 proprietary database of the Pasteur Institute of Guadeloupe, which is an updated version of the previously released SpolDB4 database (available at http://www.pasteur-guadeloupe.fr:8081/SITVIT_ONLINE/). Currently, SITVIT2 contains 90,000 isolates. In this database, spoligotype international type (SIT) designates a spoligotyping pattern shared by two or more patient isolates.

2.4. MLST

Fourteen genes distributed around the chromosome of *M. tuberculosis* H37Rv (accession number NC_000962.2) were initially chosen for MLST scheme. These housekeeping genes were chosen based on their putative function, on their use in MLST schemes for other bacterial species (http://pubmlst.org/databases.shtml) and on previous studies on genetic polymorphisms in *M. tuberculosis* (Baker et al., 2004; Gutierrez et al., 2005; Rouse et al., 1995). Primers were designed based on the nucleotide sequences of the *M. tuberculosis* H37Rv deposited in GenBank (Accession No: NC_000962.2), using PrimerQuest software (Integrated DNA Technologies, Coralville, Iowa), to amplify 465 to 787 bp of the genes described in Table 1.

The gene fragments were amplified from genomic DNA with the primers described in Table 1, under the following conditions. Amplifications were carried out in a total volume of 50 μL with 60 ng of template DNA, 0.25 pmol of primers (Invitrogen, Carlsbad, CA), 5 μL of 10× buffer, 15 mmol L^{-1} $MgCl_2$ (Invitrogen, Carlsbad, CA), 4 μL of 2 mmol L^{-1} deoxynucleoside triphosphate mix (Eppendorf, New York, USA), and 0.125 U Taq High Fidelity (Invitrogen, Carlsbad, CA). PCR reaction was performed in a Mastercycler Gradient (Eppendorf) for all genes as it follows: initial denaturation 94 °C for 2 min, 30 cycles at 94 °C for 1 min, annealing temperature 60 °C for 1 min, 72 °C for 1 min and an additional extension at 72 °C for 5 min. *M. tuberculosis* H37Rv (ATCC 27294) was used as positive control, and reactions without DNA were used as negative control for all PCR assays. The amplified products were resolved by 1.5% agarose gel electrophoresis into bands, which were stained with ethidium bromide (0.5 μg mL^{-1}) and revealed under UV light. Expected PCR product lengths for each primer set are listed in Table 1.

PCR products were purified for sequencing with GFX PCR (GE Healthcare, Buckinghamshire, UK). Automated DNA sequencing was performed using MegaBACE 1000 DNA sequencers (GE Healthcare, Buckinghamshire, UK), with the same PCR primer sets. Each forward and reverse strand was sequenced at least twice. Raw sequences were reviewed by visual inspection using ChromasPro version 1.33 software (Technelysium Pty. Ltd). Different sequences of a given locus were given allele numbers, and each unique combination of alleles (the allelic profile) was assigned a sequence type (ST).

After a pilot study, the seven best housekeeping genes were chosen among the fourteen genes previously mentioned for MLST scheme, based on criteria that included diversity of obtained sequences, facility

Table 1
Primers used for PCR amplification and sequencing of *Mycobacterium tuberculosis* housekeeping genes for MLST analysis. The seven housekeeping genes chosen for this study are highlighted in gray background.

Gene	Protein products	Primers sequences (forward and reverse, 5′–3′)	Product (bp)
aroE	Shikimate 5-dehydrogenase	GACCTATGAGCGCATCGAATG CTGATGCAGCAACATCTGCAG	622
ddl	D-alanylalanine synthetase	GGAGCTTCCTCAGGTCAAATCA CGTGTTGATCTCGTTGATCACC	787
dnaN	DNA polymerase III beta subunit	TTGCCTAACAAGCCCGTAGACGTT TTCGGGAACTCGGCATCAAGAAGT	503
fumC	Fumarate hydratase	CAACGACGACGTGAACATGTC GGGATGTAGACGTTGAGTTCG	702
gyrA	DNA gyrase A subunit	GTCGGCATGGCAACCAATATC CAGAATGTGGGCTCGCTCGTT	628
gyrB	DNA gyrase B subunit	CAAGAGATGGCGTTCCTCAAC GACACAGCCTTGTTCACAACG	653
katG	Catalase–peroxidase–peroxynitrase	GGAATCGATGGGCTTCAAGAC CATGTCTCGGTGGATCAGCTT	742
ptA	Phosphate acétyltransférase	ATTCCCATCCCTACGGTCACTACA TATGCCCGGAACGGTCTTGATGAT	465
purA	Adenylosuccinate dehydrogenase	GCTATCGACAAGGTCACCGAG GAAGTAGTCGGTGATGCCGTT	651
purH	Phoshoribosylaminoimiazolcarboxylase ATPase subunit	TGGCGTCGTTAGCGTTTCAGCATA ATGGTGCTCACATACTCGGCCATT	472
recA	Recombinase A	CGCAAGCCTATTCATGTCGTG GAGTTGTTCAGAGGTCGTCGT	657
rpoB	*DNA-directed RNA polymerase beta subunit*	*GCTGAGCCAATTCATGGACCA* *GCATCACAGTGATGTAGTCGG*	680
sodA	Superoxide dismutase	CGAATACACCTTGCCAGACCT CCTTGGTCTGCGAGGTCGCGG	602
trpE	Anthranilate synthase component I	ATGGTCGGTTTCTTCGCCTATGAC AAGTCCACTGCACCATCACTATTCGG	512

and reproducibility amplification of the PCR products, and efficacy of sequencing using the same primers used for amplification. Thus, the following seven loci were chosen for the MLST scheme: *gyrA*, *gyrB*, *katG*, *purA*, *recA*, *rpoB* and *sodA*.

For each isolate, the different sequences present in each housekeeping gene are assigned as distinct alleles, which received an arbitrary allele number for identification. Each novel set of seven alleles was defined as sequence typing (ST), which also received an arbitrary number for identification.

2.5. Construction of dendrograms

Individual data from spoligotyping, MIRU-VNTR and MLST and a combination of the three data sets were analyzed with the software program BioNumerics 4.5 (Applied Maths®, Sint Martens Latem, Belgium), using Euclidean Distance to calculate the similarity matrix. A dendrogram based on this matrix was constructed by the UPGMA (unweighted pair group method) method. For MLST, all seven housekeeping gene sequences referring to each isolate were previously concatenated using the program ChromasPro version 1.33 (Technelysium Pty Ltd).

Isolates that showed 100% identity by the different methods were considered clusters and, based on this clustering, a table was constructed to organize the obtained results for the isolates studied by each technique (Table 2).

2.6. Discrimination index (DI)

The individual discriminatory ability of spoligotyping, MIRU and MLST was evaluated by Simpson's index of diversity, as described by Hunter and Gaston (1998). This index was also calculated for different combinations among these methods. The authors recommend that a desirable index of discrimination for a molecular typing methodology must be greater than 0.90 in order to be interpreted with confidence (maximum is 1). The best discriminatory method is the one that presents the largest index value when compared to others. This index value is commonly referred as Hunter–Gaston discriminatory index (HGDI).

3. Results and discussion

Forty-four *M. tuberculosis* isolates, after phenotypic analyses, were submitted to molecular identification and all isolates have been confirmed as *M. tuberculosis* complex. Afterwards, the isolates were genotyped using spoligotyping, MIRU-VNTR and MLST techniques (Table 2).

The spoligotyping technique showed 27 different profiles, being 56.8% (25/44) of the isolates grouped in eight clusters (A, B, C, D, E, F, G and H) (Fig. 1A). The Latin American Mediterranean (LAM) clade was the most incident with 38.6% (17/44) of the isolates, followed by T clade with 29.5% (13/44). Other clades were also present, such as Beijing, H and U. Four profiles were classified as orphans because they have not been described in the data bank SITVIT2 yet (Fig. 1A and Table 2). Earlier studies have demonstrated the predominance of these clades in Brazil (Noguti et al., 2012; Silva et al., 2009). The present results corroborated with those reported by Mendes et al. (2011) which describe a predominance of the clades LAM and T in *M. tuberculosis* Brazilian isolates.

Regarding the LAM clade, LAM9 (SIT42) was identified in 47% of the isolates (8/17) (cluster D) followed by LAM9 (SIT1075), which was present in 17.6% (3/17) (cluster B) and LAM9 (SIT1800) (cluster C), in 11.7% (2/17). Other LAM profiles (LAM1, LAM2, LAM3 and LAM6) were represented by only one isolate each. Concerning geographic distribution, the LAM clade has a ubiquitous distribution with highly predominance in America. In Brazil, SIT42 has a good adaptability and transmissibility among Brazilian patients. Among eight isolates classified in SIT42, seven were from Dourados city (Mato Grosso do Sul state) and only one from São Paulo city (São Paulo state).

Thirteen isolates were identified as T1 clade with 10 different SITs (1284, 1166, 1214, 102, 174, 291, 291, 86, 244, 53) and two isolates as Beijing clade with the same SIT1 (cluster A). These clades have ubiquitous distribution but, in this study, the Beijing clade was only present in São Paulo city. This fact may be attributed to the large immigration to this city (Table 2).

Three strains were classified as unknown (U) clade with distinct SITs (29, 560, 450). The H clade was present in five isolates distributed in two groups (F and G), with the SITs 47 (H1) and 50 (H3), respectively. The H clade was only present in São Paulo state. Two isolates (3 and 102) with different and single spoligotyping profiles were designated as orphans by the SITVIT2 database (Table 2; Figs. 1A and 2). It is important to mention that some spoligotypes found in this study were recently described for the first time in Brazil (SITs: 86, 560, 1214 and 1266), which suggests these spoligotypes may be becoming more frequent in this country (Mendes et al., 2011).

Regarding MIRU-VNTR, the same 44 isolates were analyzed using this technique and 34 different profiles could be observed. 31.8% (14/44) of the isolates were distributed in five clusters with 100% of identity (A′, B′, C′, D′ and E′) (Fig. 1B). Four clusters were formed by two isolates each (A′, B′, C′, E′). One cluster (D′) was formed by six isolates of patients from Dourados city (Table 2, Fig. 1B).

In the clusters C′, D′ and E′, it is possible to observe in Table 2 that some strains exhibit exactly the same profile typing for all three methods and were classified as real clusters. Three real clusters were related to isolates (pairs of isolates 104 and 107, 106 and 112, 113 and 117) of patients from São Paulo city and one of them was related to isolates (7 and 8) of patients from Araraquara city. These real clusters indicate that the isolates could have been recently transmitted among

Table 2
Summarized results of all 44 isolates grouped by MIRU-VNTR technique showing five clusters and the MLST results.

Isolate	City/State	Spoligotyping	Clade	SIT	Profile	MIRU-VNTR	Cluster	MLST (ST)
109	Dourados/MS	777770607760771	LAM9	1075		122326153225	Unclustered	1
125	São Paulo/SP	777777737760771	T1	86		123326143204	Unclustered	1
124	São Paulo/SP	777777477720771	H3	No correlation		124301152314	Unclustered	3
111	Dourados/MS	777777407760771	LAM9	1800		124326133226	A'	3
031-307	Campo Grande/MS	777777607760771	LAM9	42		124326133226		1
5	Araraquara/SP	777777607560771	LAM6	64		124326133324	Unclustered	1
102	Dourados/MS	777777660010371	Orphan	Orphan		124326153224	Unclustered	1
114	Dourados/MS	777777037760771	T1	174		124327133224	Unclustered	1
110	Dourados/MS	777770607760771	LAM9	1075		124327133225	Unclustered	1
123	São Paulo/SP	777000000000371	U	560		124425043111	Unclustered	1
15CF	São Paulo/SP	777617777760771	T1	1214		222225052311	Unclustered	2
07CF	São Paulo/SP	000000000003771	Beijing	1		223126152331	Unclustered	1
100	Dourados/MS	741703777760771	T1	1284		223226163221	Unclustered	1
121	São Paulo/SP	777777777760771	T1	53		223325163214	Unclustered	4
126	São Paulo/SP	000000000003771	Beijing	1		223325183503	Unclustered	4
2	Araraquara/SP	777777777760700	T1	51		223326143323	Unclustered	1
1	Araraquara/SP	777777637760771	T1	2070		224015163321	Unclustered	1
115	Dourados/MS	777777737760771	T1	86		224124132321	Unclustered	5
10	Araraquara/SP	777777774020771	H1	47		224126142321	Unclustered	1
3	Araraquara/SP	777737607760700	Orphan	Orphan		224126152321	B'	1
9	Araraquara/SP	777777774020771	H1	47		224126152321		1
104	Dourados/MS	777777607760771	LAM9	42		224225163321	C'	1
107	Dourados/MS	777777607760771	LAM9	42		224225163321		1
101	Dourados/MS	777777607760771	LAM9	42		224226142321	Unclustered	1
108	Dourados/MS	777770607760771	LAM9	1075		224226153321	D'	1
103	Dourados/MS	777777407760771	LAM9	1800		224226153321		1
106	Dourados/MS	777777607760771	LAM9	42		224226153321	D'	1
112	Dourados/MS	777777607760771	LAM9	42		224226153321		1
113	Dourados/MS	777777777760771	T1	53		224226153321	D'	1
117	Dourados/MS	777777777760771	T1	53		224226153321		1
105	Dourados/MS	777777607760771	LAM9	42		224226153322	Unclustered	1
031-1235	Campo Grande/MS	776137607760771	LAM3	211		224226153323	Unclustered	1
6	Araraquara/SP	677777607760771	LAM1	20		224226163321	Unclustered	1
116	Dourados/MS	777777677760771	T1	291		224227153321	Unclustered	1
122	São Paulo/SP	777777607760771	LAM9	42		224325131311	Unclustered	3
129	São Paulo/SP	677737607760771	LAM2	17		224325143211	Unclustered	1
127	São Paulo/SP	777703777760771	T1	102		224325153312	Unclustered	1
4	Araraquara/SP	760001400000171	U	29		224326133324	Unclustered	1
120	São Paulo/SP	777777774020771	H1	47		225313153314	Unclustered	1
7	Araraquara/SP	777777777720771	H3	50		225314153323	E'	1
8	Araraquara/SP	777777777720771	H3	50		225314153323		1
118	Dourados/MS	777776770000000	U	450		225325153224	Unclustered	1
119	Dourados/MS	777377777760771	T1	1166		233325143324	Unclustered	1
128	São Paulo/SP	777777777760601	T1	244		243260422113	Unclustered	6

Note: Real clusters are highlighted with a gray background.

the patients, since all of them presented the same profile and resided in the same cities (Table 2), considering the three studied methods.

Multilocus sequence typing (MLST) is broadly and successfully used for a variety of bacteria and the number of new databases available to study different bacterial species is constantly growing. Currently, there are approximately ninety different databases for microorganisms, including fungi and bacteria (http://pubmlst.org/databases.shtml). Although some authors suggest the possibility that MLST is not applicable for typing isolates of *M. tuberculosis* due to the restricted structural polymorphism of gene sequences, this approach has still not been tested specifically for this species (Comas et al., 2009; Supply et al., 2002). On the other hand, there are examples of successful application of MLST for nontuberculous mycobacteria (NTM), such as *Mycobacterium abscessus*, *Mycobacterium ulcerans*, *Mycobacterium marinum* and others (Macheras et al., 2011; Mignard and Flandrois, 2008; Stinear et al., 2000). Therefore, in this study, the authors propose to standardize and test the MLST technique for *M. tuberculosis* and to compare the discriminatory power of this technique with MIRU-VNTR and spoligotyping.

For MLST analysis, the fourteen housekeeping gene fragments were amplified and sequenced for all 44 *M. tuberculosis* isolates. The sequences obtained for each one of the seven loci referring to each isolate were individually compared to all other isolates, and the alleles were named consecutively after numbers. Many gene fragments

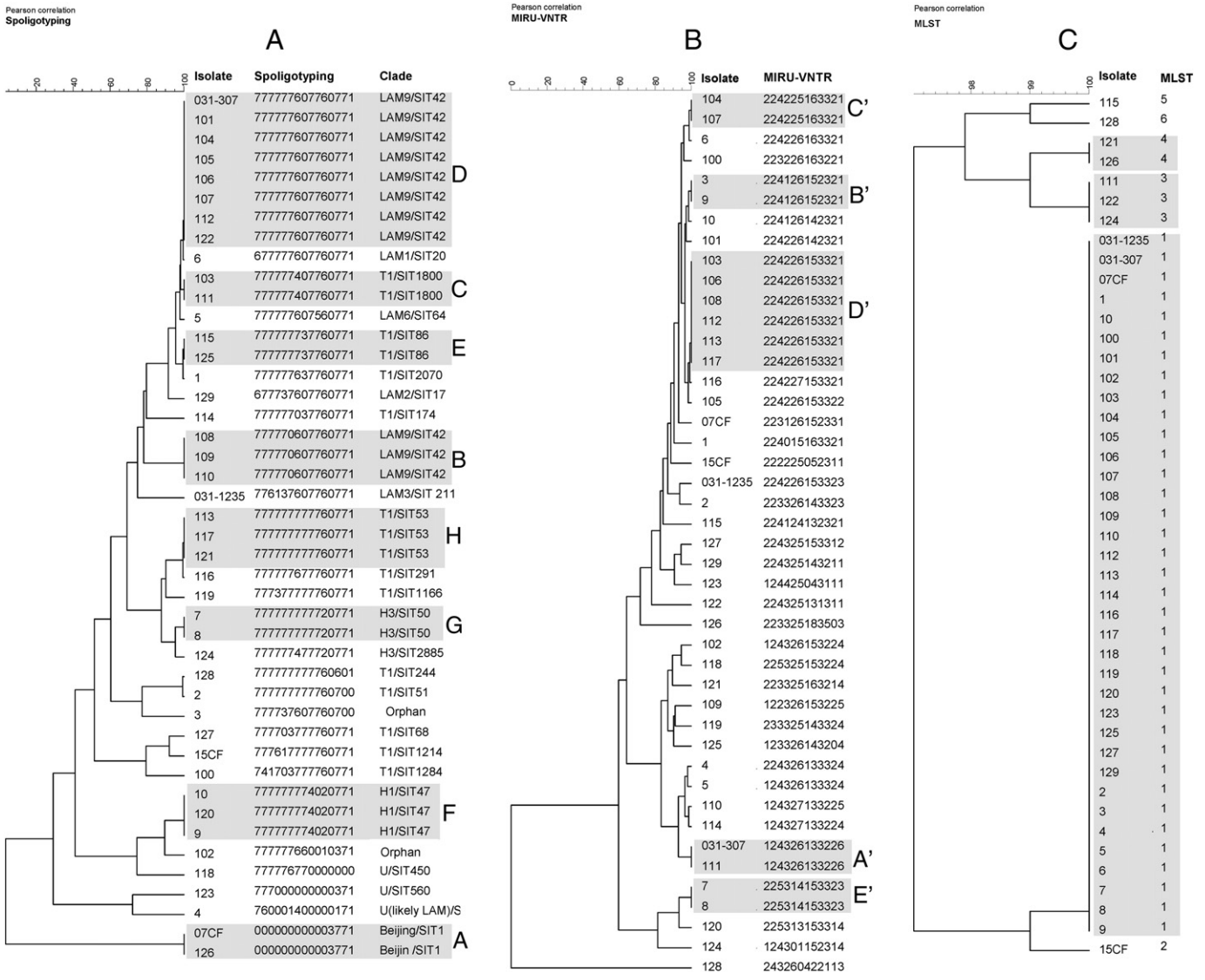
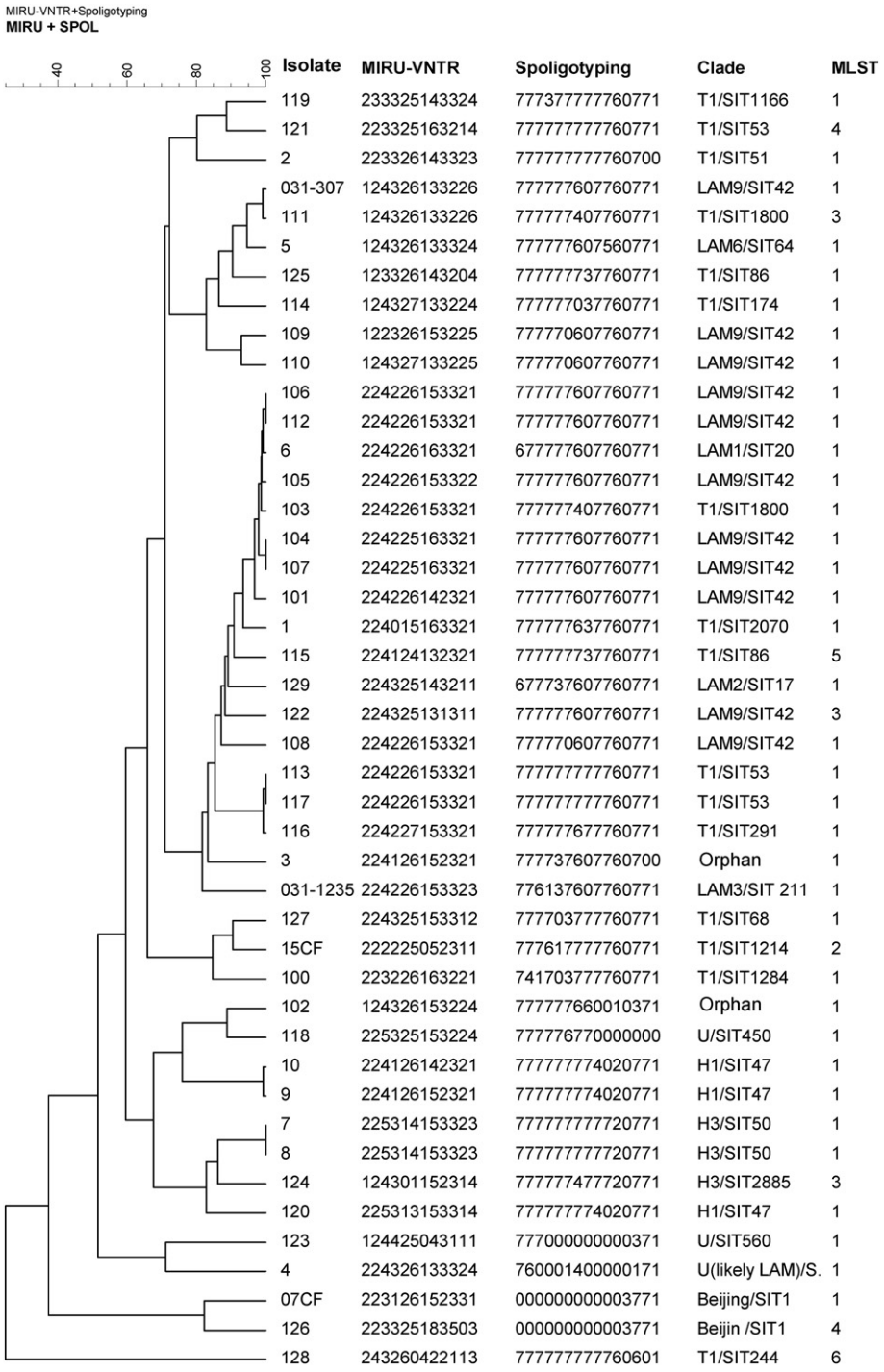


Fig. 1. Dendrograms showing the genetic relationships among 44 strains of *M. tuberculosis* based on spoligotyping (A), MIRU-VNTR (B) and MLST (C). The dendrogram was generated using the Pearson correlation and the unweighted pair group method (UPGMA). Clusters of identical patterns are indicated in gray boxes and represented by different letters. MIRU = mycobacterial interspersed repetitive units; VNTR = variable number of tandem repeats; and MLST = multilocus sequence typing.

analyzed after sequencing had identified a single allele for all strains (*aroE*, *ddl*, *dnaN*, *fumC*, *pta*, *purA*, *purH*, *recA*, *rpoB*, *trpE*). Based on the criteria described in the methodology, the seven best were chosen to be used in the MLST scheme, namely *gyrA*, *gyrB*, *katG*, *purA*, *recA*, *rpoB*, and *sodA*. Among the seven housekeeping genes analyzed, the loci that presented major genetic variability were *sodA*, showing three different alleles, followed by *gyrA*, *gyrB* and *katG*, with two alleles each. The loci *purA*, *recA* and *rpoB* showed a single allele for all isolates (data not shown). The allelic profile of each isolate generated a total of six different sequence types (STs). The most frequent ST was ST1, comprehending 36 isolates (81.8%), followed by ST3 with 3 isolates (6.8%), ST4 with 2 isolates (4.5%) and STs 2, 5 and 6 comprising one isolate each (2.3%) (Fig. 1C and Table 2). The dendrogram represented in Fig. 1C shows the genetic relationship among the studied isolates and in which it is possible to observe the separation of clusters according to the STs obtained for each isolate.

Based on the dendrograms obtained for each typing method applied to the isolates of this study, the Simpson's index of diversity was calculated to measure the discriminatory power of MIRU-VNTR, spoligotyping and MLST and the obtained values were 0.98, 0.96 and 0.33, respectively. Comparing these obtained values, it is clear that, according to the theory of Hunter and Gaston (1998), the techniques MIRU-VNTR and spoligotyping have a much higher discriminatory power than MLST. Moreover, although they have obtained similar values, MIRU-VNTR has greater discriminatory power in comparison to spoligotyping. The superiority of MIRU-VNTR has been previously described in former studies, which found rates of HGDI similar to those found in the present study (Dong et al., 2010; Mulenga et al., 2010). In contrast, MLST showed a very low value to be interpreted with confidence. The difference of discriminatory power among the three methods is easily perceived by the topology of the three dendrograms presented in Fig. 1, in which it is possible to observe only three clusters in MLST and that the great majority of isolates are grouped into only one cluster (Fig. 1C). In contrast, MIRU and spoligotyping presented a higher number of clusters with less isolates each (Fig. 1A and B).

As expected, the combinatorial analysis of MIRU-VNTR and spoligotyping techniques with MLST showed a significant reduction of discriminatory power in both cases. However, the combination of MIRU-VNTR with spoligotyping demonstrated a significant increase in the ability to discriminate isolates of the study. While MIRU-VNTR and spoligotyping presented HGDI of 0.98 and 0.96, respectively, both



Isolate	MIRU-VNTR	Spoligotyping	Clade	MLST
119	233325143324	777377777760771	T1/SIT1166	1
121	223325163214	777777777760771	T1/SIT53	4
2	223326143323	777777777760700	T1/SIT51	1
031-307	124326133226	777777607760771	LAM9/SIT42	1
111	124326133226	777777407760771	T1/SIT1800	3
5	124326133324	777777607560771	LAM6/SIT64	1
125	123326143204	777777737760771	T1/SIT86	1
114	124327133224	777777037760771	T1/SIT174	1
109	122326153225	777770607760771	LAM9/SIT42	1
110	124327133225	777770607760771	LAM9/SIT42	1
106	224226153321	777777607760771	LAM9/SIT42	1
112	224226153321	777777607760771	LAM9/SIT42	1
6	224226163321	677777607760771	LAM1/SIT20	1
105	224226153322	777777607760771	LAM9/SIT42	1
103	224226153321	777777407760771	T1/SIT1800	1
104	224225163321	777777607760771	LAM9/SIT42	1
107	224225163321	777777607760771	LAM9/SIT42	1
101	224226142321	777777607760771	LAM9/SIT42	1
1	224015163321	777777637760771	T1/SIT2070	1
115	224124132321	777777737760771	T1/SIT86	5
129	224325143211	677737607760771	LAM2/SIT17	1
122	224325131311	777777607760771	LAM9/SIT42	3
108	224226153321	777770607760771	LAM9/SIT42	1
113	224226153321	777777777760771	T1/SIT53	1
117	224226153321	777777777760771	T1/SIT53	1
116	224227153321	777777677760771	T1/SIT291	1
3	224126152321	777737607760700	Orphan	1
031-1235	224226153323	776137607760771	LAM3/SIT 211	1
127	224325153312	777703777760771	T1/SIT68	1
15CF	222225052311	777617777760771	T1/SIT1214	2
100	223226163221	741703777760771	T1/SIT1284	1
102	124326153224	777777660010371	Orphan	1
118	225325153224	777776770000000	U/SIT450	1
10	224126142321	777777774020771	H1/SIT47	1
9	224126152321	777777774020771	H1/SIT47	1
7	225314153323	777777777720771	H3/SIT50	1
8	225314153323	777777777720771	H3/SIT50	1
124	124301152314	777777477720771	H3/SIT2885	3
120	225313153314	777777774020771	H1/SIT47	1
123	124425043111	777000000000371	U/SIT560	1
4	224326133324	760001400000171	U(likely LAM)/S.	1
07CF	223126152331	000000000003771	Beijing/SIT1	1
126	223325183503	000000000003771	Beijin /SIT1	4
128	243260422113	777777777760601	T1/SIT244	6

Fig. 2. Dendrogram constructed based on the data association of spoligotyping and MIRU-VNTR from the 44 studied *M. tuberculosis* isolates. The dendrogram was generated using the Pearson correlation and the unweighted pair group method (UPGMA). MIRU = mycobacterial interspersed repetitive units; VNTR = variable number of tandem repeats; and MLST = multilocus sequence typing.

techniques associated showed an index of 0.99, which is nearly considered a desirable perfect index. The dendrogram resulting from the association of these two methods is illustrated in Fig. 2, in which it is possible to observe that some isolates that were clustered by MIRU-VNTR and spoligotyping (Fig. 1 A and B) were separated after the combination of these two methods. These results corroborate with the findings of a study conducted by Noguti et al. (2012), which also shows the increase in discriminatory power by combining these two techniques for isolates from a low-endemic setting population. Mendes et al. (2011) also demonstrated that some isolates clustered by spoligotyping in 100% identity were separated when combined with the MIRU-VNTR technique.

Although MLST is broadly used as a useful tool for molecular epidemiology of many bacterial species, the results presented here demonstrate that MLST has low discriminatory power for *M. tuberculosis* isolates, suggesting that it is not applicable for this species. This may be explained due to the genome of *M. tuberculosis* being genetically monomorphic harboring little DNA sequence diversity (Comas et al., 2009) and also due to the sequence conservation of the housekeeping genes, thereby MLST might not have sufficient discriminatory power for *M. tuberculosis* isolates. Recently, the advent of next generation sequencing (NGS) has opened new possibilities for epidemiological and phylogenetic studies, especially for genetically monomorphic

bacteria. This progress may significantly improve future studies for MLST of *M. tuberculosis*, allowing the analysis of a large number of genes in a shorter period of time, increasing the chances to distinguish isolates that have the same MLST profile when analyzed by the traditional method.

In conclusion, in this study, three molecular typing methods (MIRU-VNTR, spoligotyping and MLST) were tested for 44 *M. tuberculosis* isolates, of which MLST was fully tested for this species for the first time in the present study. Among the three methods, MLST did not show satisfactory results and is not indicated as an appropriate method to differentiate strains of *M. tuberculosis*. MIRU-VNTR showed better discriminatory power when compared to spoligotyping, however, a combination of both methods provided the greatest level of discrimination and, therefore, it is the most useful genotyping tool to be applied in *M. tuberculosis* isolates.

Funding source

This study was supported by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) (Proc. PNPD 033497/2008-86).

Acknowledgments

The authors are grateful to Wilson A. Silva, Jr., Adriana A. Marques and Cristiane A. Ferreira at the Laboratory of Molecular Genetics and Bioinformatics, Department of Genetics, Faculty of Medicine of Ribeirão Preto, University of São Paulo, for DNA sequencing and Daisy N. Sato at the Adolfo Lutz Institute, Ribeirão Preto São Paulo, Brazil, for critical reading of the manuscript.

References

Allix-Béguec, C., Fauville-Dufaux, M., Supply, P., 2008. Three-year population-based evaluation of standardized mycobacterial interspersed repetitive-unit-variable-number tandem-repeat typing of *Mycobacterium tuberculosis*. J. Clin. Microbiol. 46, 1398–1406.

Alonso-Rodríguez, N., Martínez-Lirola, M., Herránz, M., Sanchez-Benitez, M., Barroso, P., INDAL-TB group, et al., 2008. Evaluation of the new advanced 15-loci MIRU-VNTR genotyping tool in *Mycobacterium tuberculosis* molecular epidemiology studies. BMC Microbiol. 8, 34.

Baker, L., Brown, T., Maiden, M.C., Drobniewski, F., 2004. Silent nucleotide polymorphisms and a phylogeny for *Mycobacterium tuberculosis*. Emerg. Infect. Dis. 10, 1568–1577.

Comas, I., Homolka, S., Niemann, S., Gagneux, S., 2009. Genotyping of genetically monomorphic bacteria: DNA sequencing in *Mycobacterium tuberculosis* highlights the limitations of current methodologies. PLoS One 12, e7815.

Cowan, L.S., Mosher, L., Diem, L., Massey, J.P., Crawford, J.T., 2002. Variable-number tandem repeat typing of *Mycobacterium tuberculosis* isolates with low copy numbers of IS6110 by using mycobacterial interspersed repetitive units. J. Clin. Microbiol. 40, 1592–1602.

Cronin, W.A., Golub, J.E., Magder, L.S., Baruch, N.G., Lathan, M.J., Mukasa, L.N., et al., 2001. Epidemiologic usefulness of spoligotyping for secondary typing of *Mycobacterium tuberculosis* isolates with low copy numbers of IS6110. J. Clin. Microbiol. 39, 3709–3711.

Dong, H., Shi, L., Zhao, X., Sang, B., Lv, B., Liu, Z., et al., 2010. Diversity of *Mycobacterium tuberculosis* genotypes circulating in Ndola, Zambia. BMC Infect. Dis. 17, 10–177.

Gutierrez, C., Brisse, S., Brosch, R., Fabre, M., Omais, B., Marmiesse, M., et al., 2005. Ancient origin and gene mosaicism of the progenitor of *Mycobacterium tuberculosis*. PLoS Pathog. 1, 1–7.

Hunter, P.R., Gaston, M.A., 1998. Numerical index of the discriminatory ability of typing systems: an application of Simpson's index of diversity. J. Clin. Microbiol. 26, 2465–2466.

Kamerbeek, J., Schouls, L., Kolk, A., van Agterveld, M., van Soolingen, D., Kuijper, S., et al., 1997. Simultaneous detection and strain differentiation of *Mycobacterium tuberculosis* for diagnosis and epidemiology. J. Clin. Microbiol. 35, 907–914.

Leão, S.C., Martin, A., Megia, G.I., Palomino, J.C., Robledo, J., Telles, M., et al., 2004. Pratical handbook for the phenotypic and genotypic identification of mycobacteria. Vanden Broelle, Brugges (164 pp. [URL: http://www.esmycobacteriology.eu/Inco.htm]).

Lukinmaa, S., Nakari, U.M., Eklund, M., Siitonen, A., 2004. Application of molecular genetic methods in diagnostics and epidemiology of food-borne bacterial pathogens. APMIS 112, 908–929.

Macheras, E., Roux, A.L., Bastian, S., Leão, S.C., Palaci, M., Sivadon-Tardy, V., et al., 2011. Multilocus sequence analysis and rpoB sequencing of *Mycobacterium abscessus* (sensu lato) strains. J. Clin. Microbiol. 49, 491–499.

Maiden, M.C., 2006. Multilocus sequence typing of bacteria. Annu. Rev. Microbiol. 60, 561–588.

Maiden, M.C., Bygraves, J.A., Feil, E., Morelli, G., Russell, J.E., Urwin, R., et al., 1998. Multilocus sequence typing: a portable approach to the identification of clones within populations of pathogenic microorganisms. Proc. Natl. Acad. Sci. U. S. A. 95, 3140–3145.

Martens, M., Dawyndt, P., Coopman, R., Gillis, M., De Vos, P., Willems, A., 2008. Advantages of multilocus sequence analysis for taxonomic studies: a case study using 10 housekeeping genes in the genus *Ensifer* (including former *Sinorhizobium*). Int. J. Syst. Evol. Microbiol. 58, 200–214.

Mathema, B., Kurepina, N.E., Bifani, P.J., Kreiswirth, B.N., 2006. Molecular epidemiology of tuberculosis: current insights. Clin. Microbiol. Rev. 19, 658–685.

Mazars, E., Lesjean, S., Banuls, A.L., Gilbert, M., Vincent, V., Gicquel, B., et al., 2001. High-resolution minisatellite-based typing as a portable approach to global analysis of *Mycobacterium tuberculosis* molecular epidemiology. PNAS 98, 1901–1906.

Mendes, N.H., Melo, F.A., Santos, A.C., Pandolfi, J.R., Almeida, E.A., Cardoso, R.F., et al., 2011. Characterization of the genetic diversity of *Mycobacterium tuberculosis* in São Paulo city, Brazil. BMC Res. Notes 4, 269.

Mignard, S., Flandrois, J.P., 2008. A seven-gene, multilocus, genus-wide approach to the phylogeny of mycobacteria using supertrees. Int. J. Syst. Evol. Microbiol. 58, 1432–1441.

Mulenga, C., Shamputa, I.C., Mwakazanga, D., Kapata, N., Portaels, F., Rigouts, L., 2010. Diversity of *Mycobacterium tuberculosis* genotypes circulating in Ndola, Zambia. BMC Infect. Dis. 17, 10–177.

Noguti, E.N., Leite, C.Q., Malaspina, A.C., Santos, A.C., Hirata, R.D., Hirata, M.H., et al., 2012. Genotyping of *Mycobacterium tuberculosis* isolates from a low-endemic setting in northwestern state of Paraná in Southern Brazil. Mem. Inst. Oswaldo Cruz 105, 779–785.

Rouse, D.A., Li, Z., Bai, G.H., Morris, S.L., 1995. Characterization of the *katG* and *inhA* genes of isoniazidresistant clinical isolates of *Mycobacterium tuberculosis*. Antimicrob. Agents Chemother. 39, 2472–2477.

Sequera, M.C., Delgado, V.S., Araque, M.W., Torrealba, M.O., Núñes, M.R., Mata, O.J., et al., 2008. *Mycobacterium tuberculosis*: Espoligotipos em el estados Carabobo, Venezuela. Rev. Chil. Infectol. 25, 362–367.

Silva, A.B.S., Groll, A.V., Feliz, C., Conceição, F.R., Spies, F.S., Scaini, C.J., et al., 2009. Clonal diversity of *Mycobacterium tuberculosis* isolated in a sea port city in Brazil. Tuberculosis 89, 443–447.

Sola, C., Filliol, I., Gutirrez, M.C., Mokrousov, I., Vincent, V., Rastogi, N., 2001. Spoligotype database of *Mycobacterium tuberculosis*: biogeografic distribution of shared types and epidemiologic and phylogenetic perspectives. Emerg. Infect. Dis. 7, 390–396.

Stackebrandt, E., Frederiksen, W., Garrity, G.M., Grimont, P.A., Kämpfer, P., Maiden, M.C.J., et al., 2002. Report of the ad hoc committee for the reevaluation of the species definition in bacteriology. Int. J. Syst. Evol. Microbiol. 52, 1043–1047.

Stinear, T.P., Jenkin, G.A., Johnson, P.D., Davies, J.K., 2000. Comparative genetic analysis of *Mycobacterium ulcerans* and *Mycobacterium marinum* reveals evidence of recent divergence. J. Bacteriol. 182, 6322–6330.

Supply, P., Allix, C., Lesjean, S., Cardoso-Oelemann, M., Rüsch-Gerdes, S., Willery, E., et al., 2006. Proposal for standardization of optimized mycobacterial interspersed repetitive unit-variable-number tandem repeat typing of *Mycobacterium tuberculosis*. J. Clin. Microbiol. 44, 4498–4510.

Supply, P., Lesjean, S., Savine, E., Kremer, K., van Soolingen, D., Locht, C., 2002. Automated high-throughput genotyping for study of global epidemiology of *Mycobacterium tuberculosis* based on mycobacterial interspersed repetitive units. J. Clin. Microbiol. 39, 3563–3571.

Supply, P., Mazars, E., Lesjean, S., Vincent, V., Gicquel, B., Locht, C., 2000. Variable human minisatellite-like regions in the *Mycobacterium tuberculosis* genome. Mol. Microbiol. 36, 762–771.

van Embden, J.D., Cave, M.D., Crawford, J.T., Dale, J.W., Eisenach, K.D., Gicquel, B., et al., 1993. Strain identification of *Mycobacterium tuberculosis* by DNA fingerprinting: recommendations for a standardized methodology. J. Clin. Microbiol. 31, 406–409.

World Health Organization, 2012. Global Tuberculosis Report 2012. WHO Press, Geneva, Switzerland, pp. 1–9.

Zaman, K., 2010. Tuberculosis: a global health problem. Editorial. J. Health Popul. Nutr. 28, 111–113.